working paper
department
of economics

RATIONAL DEBATE AND
ONE-DIMENSIONAL CONFLICT

David Spector

massachusetts
institute of
technology

50 memorial drive
cambridge, mass. 02139

WORKING PAPER
DEPARTMENT
OF ECONOMICS

RATIONAL DEBATE AND
ONE-DIMENSIONAL CONFLICT

David Spector

MASSACHUSETTS
INSTITUTE OF
TECHNOLOGY

50 MEMORIAL DRIVE
CAMBRIDGE, MASS. 02142

Rational debate and one-dimensional conflict[1]

David Spector (MIT)

Revised version: February 1999

"It's like G.H. Hardy's old crack: If the Archbishop of Canterbury says he believes in God, that's all the way of business, but if he says he doesn't, one can take it he means what he says."

C.P. Snow, *The Affair*

Abstract

This paper studies repeated communication regarding a multidimensional collective decision in a large population. When preferences coincide but beliefs about the consequences of the various decisions diverge, it is shown, under some specific assumptions, that public communication causes the disagreement between beliefs either to vanish or to become one-dimensional at the limit. Multidimensional disagreement allows indeed for many directions of communication, including some which are orthogonal to the conflict, along which agents can communicate credibly. The possible convergence towards a one-dimensional conflict where no further communication takes place may be related to the empirically observed geometry of the political conflict in many countries.

1. Introduction

This paper develops a model of repeated communication between agents facing a collective decision. We consider a situation where agents' interests coincide, while their beliefs about the uncertain payoffs of the collective decision diverge, and we analyze the dynamics of beliefs induced by strategic communication regarding this uncertainty. The main result states that the disagreement cannot remain multidimensional forever, as beliefs either converge all towards the truth, or towards some one-dimensional disagreement. We model communication as a cheap-talk game[1]: in every period, a randomly selected agent observes a signal about the state of the world and talks about it publicly, which causes all agents to update their beliefs. Then, another agent is randomly selected to make the collective decision. The information structure allows to ascribe the possible persistence of some disagreement to a communication failure, as the infinite sequence of signals would be enough, if known to all agents, to let them learn the truth.

The intuition of the main result has a simple geometric expression: a one-dimensional conflict can be a steady-state, because there are only two "directions" of communication, so that a "left-wing" agent always wants to report a "left-wing" signal, and conversely, making communication impossible. The same logic implies that, as long as the disagreement is multidimensional, agents manage to communicate credibly along a more "neutral" direction, orthogonal to their conflict, and beliefs keep changing. These two arguments together imply that, starting from a multidimensional disagreement, repeated communication causes beliefs to change and converge toward a steady-state, characterized by full agreement or one-dimensional disagreement.

However general the intuition seems, its formal modeling requires very specific assumptions: agents are infinitely impatient, communication is bound to be public, there is a continuum of agents, the payoffs of the collective decision are not observed, and the model only considers a particular utility function and signal structure. Under these assumptions, we show that for a positive measure set of initial beliefs, convergence to a one-dimensional conflict is a positive probability event.

We believe this result may be of interest for two main reasons. From a theoretical viewpoint, we know relatively little about communication in games with common knowledge of diverging priors, of which this model is a case. More important in our view, our results might be related to an empirical fact generally considered a

---

[1] See Crawford and Sobel (1982).

puzzle: contrary to the predictions of the most natural economic reasoning[2], political conflicts tend be organized along a one-dimensional axis[3]. Without claiming that our model explains this, we believe that this coincidence pleads in favor of taking seriously the idea, explored in a growing literature[4], that political conflicts are as much about beliefs as about interests, so that studying communication about collective decisions is probably an important research direction for formal political science.

The paper is organized as follows: section 2 introduces what we view as the simplest possible model allowing us to make our point; section 3 analyzes its steady-states; section 4 describes its convergence and stability properties, and section 5 concludes.

2. The model

2.1. Players and preferences

We consider a continuum of agents partitioned into two groups (1 and 2), and engaged into repeated communication over a discrete infinite horizon t=0,1,... In every period t a collective decision $M_t = (x_t, y_t) \in \Re^2$ is to be taken. Agents all have the same state-contingent utility function U(M,s), depending on the decision M and some uncertain state of the world s. U is given by

$$U(M,s) = -\|M - M_s\|^2$$

where $M_s = (x_s, y_s)$ is the optimal decision when the state of the world is s. The state of the world is unknown to the agents, and remains the same in all periods.

In addition, agents have an infinite rate of time preference, so that in period t they maximize the expectation of $U(M_t)$.

Along each of the two dimensions, there are two possible ideal decisions: x_s and y_s can each be equal to either 0 or 1. There are therefore four possible optimal decisions (0,0),

---

[2] See for example Arrow (1963).
[3] Poole and Rosenthal (1991), Snyder (1996).
[4] For example, Banerjee and Somanathan (1998), Piketty (1995).

(0,1), (1,0) and (1,1). Since the state of the world is relevant only through the optimal decision, we will identify them and write (i,j) for the state of the world where the optimal decision is (i,j).

A belief μ about the state of the world is an element $\left(\mu_{00}, \mu_{10}, \mu_{01}, \mu_{11}\right)$ of the three-dimensional simplex Δ.

We assume that all agents belonging to group i have identical beliefs in period zero, denoted μ^{i0}. The difference between beliefs across groups is the only heterogeneity between agents.

We first characterize the indirect preferences induced by some belief μ. The quadratic state-contingent utility function yields the following simple relationship between beliefs and induced preferences: indifference curves corresponding to indirect preferences are circles around the most preferred decision, which is simply the expected most preferred decision.

Lemma 1: An agent's indirect preferences given a belief μ about the state of the world are represented by the indirect utility function

$$V(M) = -\left\| M - M*(\mu) \right\|^2$$

with $M*(\mu) = E\left(M_s | \mu\right) = \left(\mu_{10} + \mu_{11}, \mu_{01} + \mu_{11}\right)$

Proof: Appendix.

Remark: A belief μ can be any element of the three-dimensional simplex Δ, but as far as the induced preferences are concerned, only M*(μ), which varies in a two-dimensional set, matters. The information which is conveyed by the belief μ in addition to that conveyed by M*(μ) is related to the correlation between the agent's beliefs about x_S and about y_S: infinitely many beliefs μ correspond the same M*(μ). Although this additional information is irrelevant for the induced preferences given quadratic preferences, it is relevant for the analysis of the communication game, since it affects the way agents update new information. One might argue that since we view the various dimensions of the collective choice as "independent" (as mentioned in the introduction),

3

we should restrict our attention to beliefs that are independent across dimensions, that is such that $\mu((x_s, y_s) = (i, j)) = \mu(x_s = i)\mu(y_s = j)$. This is wrong, because with the dynamics generated by our communication game (described below), beliefs would generally cease to be independent after communication took place.

Notation: Throughout the paper we are going to write Q for the square [0,1]x[0,1]. Lemma 1 implies that all most preferred decisions belong to Q.

2.2. The communication game

The following sequence of events is infinitely repeated:

(i)   A signal is drawn randomly according to a probability distribution depending on the true state of the world.
(ii) A randomly selected agent observes the signal and everyone learns "what the signal is about" (see section 2.3 below).
(iii) The agent who observed the signal speaks.
(iv) A randomly selected agent belonging to the group other than the speaker's takes the collective decision.

To keep the repeated game simple enough, we assume that in every odd (resp. even) period, prior to decision-making, an agent (the speaker) is randomly chosen in group 1 (resp. group 2) according to the uniform distribution. The randomization is independent across periods. The speaker then observes a signal σ_t providing some information about the state of the world, updates his beliefs, and can attempt to communicate to other agents by sending a message. An agent in the other group (the receiver) is then randomly chosen to be a dictator in period t and take the collective decision. We assume that communication is public: when an agent talks, he has to talk to everybody, and not only to the agents of his own group. We are going to write s(t) and r(t) respectively for the speaker's and the receiver's group in period t, so that (s(t),r(t)) is (1,2) if t is odd, and (2,1) otherwise.

The assumptions of a continuum of agents and of public communication imply that, except for the finite number of agents who directly observed a signal, all the other received the same information, coming exclusively from publicly sent messages.

Therefore in period t almost all agents in group i (i=1,2) have the same belief μ^{it}. With probability 1, neither the agent who observes a signal in period t nor the dictator directly observed a signal in an earlier period. This implies, with the expression B(μ,α) denoting the belief held after updating the information α starting from the belief μ, that the receiver's decision following a message m_t sent by the speaker is $M*(B(\mu^{r(t)t},m_t))$. Given the form of indirect preferences and the fact that the speaker's belief when he speaks is $B(\mu^{s(t)t},\sigma_t)$, the speaker chooses m_t to minimize

$$\left\| M*(B(\mu^{s(t)t},\sigma_t)) - M*(B(\mu^{r(t)t},m_t)) \right\|$$

In other words, the speaker sends the message which minimizes the distance between his new most preferred decision and the future dictator's.

We assume that although agents act, whenever they can, in order to maximize the expected utility derived from the collective decision, they do not observe their own utility. Therefore, agents learn nothing on their own and their beliefs are affected only by the messages they hear. This strong assumption is needed if we are to focus on limits on learning imposed by communication. If agents could learn from their utility level, then they would learn the true state of the world at once: except for a zero-measure set of decisions, the four possible states of the world yield four different utility levels, so observing utility would amount to observing the state of the world[5].

2.3. The signal structure

To make the problem interesting, signals must be neither too uninformative, nor too informative. On the one hand, we are interested in the extent to which strategic communication may leave room for some disagreement. Therefore, a desirable feature of the model should be that making the infinite sequence of signals public would cause all agents to learn the true state of the world (under the mild assumption that initial beliefs assign a non-zero probability to the truth): this will imply that any failure to learn the truth results from a communication failure. On the other hand, to make strategic communication non trivial, signals should not be completely informative: if they were, then given identical preferences, the agent who observed the signal would

---

[5] A similar assumption is made by Hart (1985).

simply report it, and would be believed.

We make some very specific assumptions about the signal structure. It is assumed to be rich enough to allow not only for pure signals (providing information about x_S only or y_S only) but also for "mixed" signals providing information simultaneously along both dimensions, with various "ratios of informativeness" along both directions. We assume that this relative informativeness (the "direction" of the signal) along the two dimensions is drawn at random and becomes known to all agents, but that this piece of information alone tells nothing about the state of the world. More precisely, the signal can be decomposed in two steps.

First, a direction $\theta \in [0, \pi)$ is drawn at random, from a probability distribution independent of the state of the world with strictly positive density everywhere. A "signal intensity" $a \in [0, A]$ is also drawn at random independently of the state of the world, with $A < \dfrac{1}{2\sqrt{2}}$. The direction and the signal intensity become common knowledge, and they obviously provide no information about the state of world.

Then, the agent randomly selected to be the speaker observe a signal which can take two values $S_{a,\theta}$ or $S_{a,\theta+\pi}$. We write $f_{ij}(S_{a,\theta})$ for the probability of observing $S_{a,\theta}$ if the state of the world is (i,j).

We assume that the functions f_{ij} have the following form:

$$f_{ij}(S_{a,\theta}) = \frac{1}{2} + a(\beta_i \cos\theta + \beta_j \sin\theta)$$

with $\beta_0 = -1, \beta_1 = 1$.

This implies that $f_{ij}(S_{a,\theta}) + f_{ij}(S_{a,\theta+\pi}) = 1$.

The assumption $A < \dfrac{1}{2\sqrt{2}}$ implies that $f_{ij}(S_{a,\theta}) > 0$ for all i, j, a and θ.

This functional form has the following simple property: if μ is the belief assigning equal weights to all four states of the world (so that $M*(\mu) = \left(\dfrac{1}{2}, \dfrac{1}{2}\right)$) then

$$M*(B(\mu, S_{a,\theta})) = \left(\frac{1}{2} + a\cos\theta, \frac{1}{2} + a\sin\theta\right)$$

so that θ is the direction of the change of the agent's most preferred decision, and a is the magnitude of this change.

This example allows to make the assumption of common knowledge of the signal direction more precise. It should be interpreted in the following way: the direction θ (mod π) is some commonly known information about "what the signal is about". For the particular example spelled out above, $\theta=0$ (mod π) means that the signal provides information only about x_s and not at all about y_s, since for i=0,1, $f_{i0}(S_{a,0}) = \frac{1}{2} + a\beta_i = f_{i1}(S_{a,0})$. Similarly if $\theta=\pi/2$ (mod π) then the signal provides information only about y_s, and all other directions are mixed, and are more and more informative about y_s (relative to x_s) the closer θ is to the vertical.

The signal structure assumed above is one of many which would yield the same results: as will appear throughout the proofs of the various propositions, what is needed is only that informativeness is bounded and that in infinitely many periods, each direction θ occurs and is common knowledge with a strictly positive probability. In particular, it may seem more natural to consider a signal structure which would allow for all directions of updating in the three-dimensional simplex, since there are four states of the world. The results would still hold - as long as informativeness is bounded and the updating direction is common knowledge. Beyond that, the specific functional form above is just chosen for its clarity.

2.4. Role of the various assumptions

2.4.1. Continuum of agents

This assumption is necessary: with finitely many agents, each of them would observe infinitely many signals directly with probability one, and would therefore learn the truth even without any communication.

2.4.2. Public communication

If communication were not bound to be public, then the speaker would communicate the signal at least to his own group, since there is no prior divergence within a group. Therefore each group would learn infinitely many signals, and converge to the truth with probability 1. The necessity of this assumption, as well as that of the continuum assumption, highlights the fact that this paper may be more relevant to think about the political debate (involving many agents and characterized by a "public" character) than about other communication situations.

2.4.3. Infinite impatience

This assumption is made for simplicity. Given the complexity of modeling forward-looking behavior in this stochastic context, we do not really know to what extent it is necessary for the results. At the very least, the results should carry over, by continuity, to the case of very impatient agents.

2.4.4. Knowledge of the signal direction

This assumption is made for simplicity. Together with the fact that there are only two signals along each direction (for a given intensity), it implies that the pure equilibria of the communication games are either completely uninformative or completely informative, which facilitates the analysis.

3. Steady states

As we stressed in section 2.2 above, the existence of a continuum of agents implies that in the beginning of any period t, almost all agents in group i have the same beliefs μ^{it}, resulting from updating all the messages sent between periods 0 and t starting from μ^{i0}. This, and the assumption that the random determination of speakers and receivers is independent across periods, ensures that in the beginning of period t+1, the speaker's and the receiver's respective beliefs are $\mu^{s(t)t}$ and $\mu^{r(t)t}$ with probability one. This allows us to keep track of the entire dynamics by using $\left(\mu^{1t},\mu^{2t}\right)$ as an exhaustive state variable.

As always in pure communication games, the question of equilibrium selection arises, since there always exists, alongside any other equilibrium, a "babbling"

equilibrium where no communication takes place. However, since there are only two signals (once the uncertainty about a and θ has been resolved), there exists for each belief distribution and signal direction either the babbling equilibrium only, or the babbling equilibrium and the "communicative" equilibrium where the speaker reveals the signal[6]. Whenever such a communicative equilibrium exists, we are going to select it. This defines a stochastic process $(\mu^{1t}, \mu^{2t})_{t \geq 0}$ starting from some arbitrary initial condition (μ^{10}, μ^{20}). We first characterize its steady states, which will allow us to describe its dynamics in Section 4.

A steady-state is a pair of beliefs (μ^1, μ^2) such that no further communication is possible, i.e. such that if $(\mu^{1t}, \mu^{2t}) = (\mu^1, \mu^2)$ then $(\mu^{1t+1}, \mu^{2t+1}) = (\mu^1, \mu^2)$ with probability 1.

Notations: For i,j in $\{0,1\}^2$ let δ_{ij} denote the vertex of Δ corresponding to the probability distribution assigning probability 1 to the event $\{s=(i,j)\}$, and zero to the other three events. We are going to call the segments $[\delta_{ij}, \delta_{i'j'}]$ (where $(i,i') \neq (j,j')$) the edges of Δ. They correspond to beliefs assigning a positive probability only to the two states of the world (i,j) and (i',j'). There are six such edges, since there are four vertices.

We first show why there are many steady-states characterized by a one-dimensional conflict of beliefs, that is, such that all beliefs belong to the same edge of Δ.

The proposition below states that there exist steady state belief distributions with support in the interior of any given edge.

Proposition 1: *Consider an edge L of Δ. There exists a set of steady states $(\mu^1, \mu^2) \in L^2$ which has a non-empty interior in L^2, and therefore, a strictly positive measure.*

Proof:

---

[6] We omit mixed equilibrium, since one can easily show that for any belief distribution, they occur for a zero-measure set of signal directions and intensities.

If L=[$\delta_{ij}, \delta_{i'j'}$] (with $(i,i') \neq (j,j')$) then each belief belonging to L can be summarized by a single number p in [0,1] denoting the probability assigned to the state of the world (i,j), the probability assigned to (i',j') being then equal to 1-p.

We consider a belief distribution such that the respective beliefs of groups 1 and 2 are p and p'. We assume that there exists a communicative equilibrium when the signal direction is θ, the intensity is a and the speaker belongs to group 1. Given the signal structure, for all signal directions θ'

$$\frac{1}{\Gamma} < \frac{f_{ij}(S_{a,\theta'})}{f_{i'j'}(S_{a,\theta'})} < \Gamma \quad \text{with} \quad \Gamma = \frac{1 + 2A\sqrt{2}}{1 - 2A\sqrt{2}}.$$

Therefore, if in equilibrium the speaker reports the signal, the posterior beliefs of members of groups 1 and 2 belong, respectively, to the intervals

$$\left[\frac{p}{p + (1-p)\Gamma}, \frac{p\Gamma}{p\Gamma + (1-p)} \right] \quad \text{and} \quad \left[\frac{p'}{p' + (1-p')\Gamma}, \frac{p'\Gamma}{p'\Gamma + 1 - p'} \right].$$

If $\dfrac{p\Gamma}{p\Gamma + 1 - p} < \dfrac{p'}{p' + (1-p')\Gamma}$ then whatever the signal was, the speaker's posterior belief

assigns a lower probability to (i,j) than the receiver would after observing any signal. Therefore, if the receiver expects the speaker to report the signal truthfully, then the best response for the speaker is to report the signal increasing the probability assigned by the receiver to (i,j), that is, $S_{a,\theta}$ if $\dfrac{f_{ij}(S_{a,\theta})}{f_{i'j'}(S_{a,\theta})} > \dfrac{f_{ij}(S_{a,\theta+\pi})}{f_{i'j'}(S_{a,\theta+\pi})}$, $S_{a,\theta+\pi}$ otherwise.

Since this best response is independent of which signal the speaker observed, truthful report of the signal is not an equilibrium and there is no communicative equilibrium.

By the same argument, there exists no communicative equilibrium with the speaker in group 2 if $\dfrac{p\Gamma}{p\Gamma + 1 - p} < \dfrac{p'}{p' + (1-p')\Gamma}$.

This inequality is satisfied by a positive measure subset of L^2, which yields the result.

q.e.d.

Proposition 1 is very intuitive: if all beliefs are on the same edge of Δ, the disagreement is one-dimensional. Therefore any new information (signals or messages)

changes beliefs in a way that makes new most preferred policies still belong to the same one dimensional set. If beliefs are initially far apart, then, since the informativeness of signals is limited, extreme beliefs remain extreme whatever the signals (and, therefore, the messages) they observe: posterior beliefs are determined mostly by prior beliefs rather than by new information. This implies that if the speaker is a "left-winger" and the receiver a "right-winger", the speaker would always prefer the receiver to believe he observed a "left-wing" signal, irrespective of the signal he truly observed. This means he cannot be credible, and agents are stuck in this polarized belief distribution where no further communication can take place.

The distance between priors in our model is the equivalent of a distance between preferences, and Proposition 1 is the equivalent, in our context, of results stating that the greater the distance of preferences, the lesser communication takes place[7].

Notice that these steady state beliefs distributions may or may not display correlation between x and y: if the edge containing them all is $[\delta_{0i}, \delta_{1i}]$ (resp. $[\delta_{i0}, \delta_{i1}]$) for $i=0$ or 1, then all beliefs agree that $y_S=i$ (resp. $x_S=i$) so that the disagreement is only about the dimension x (resp. y) and beliefs are not correlated. But if the edge is $[\delta_{00}, \delta_{11}]$ (resp. $[\delta_{01}, \delta_{10}]$), there is a positive (resp. negative) correlation between x_S and y_S.

What are the other steady-states? Proposition 2 below provides a partial answer by showing that a belief distribution such that no belief belongs to an edge of Δ is not a steady-state. It is the most important result of the paper, and it relies on the orthogonal argument mentioned in the introduction: as long as agents assign a positive probability to at least three states of the world, there exists a signal direction moving the receiver's induced most preferred decision "orthogonally" to the disagreement, and information can be credibly transmitted along this direction. Given the infinite horizon, communication will take place in the future with probability one, and such a belief distribution is not a steady-state. In other words, Proposition 2 establishes a partial converse of the result in Proposition 1: the steady states of Proposition 1 are almost the only ones (there also exist steady states such that all beliefs are located on different edges of Δ but we do not mention them because, as will appear below, the probability of converging toward them is zero if the initial beliefs are not degenerate).

Before proving Proposition 2, we need to state a few general results about how a signal

---

[7] For example, Crawford and Sobel (1982).

affects the most preferred decision of an agent whose belief does not belong to an edge of Δ. Lemma 2 below states a few simple properties about the change of the most preferred decision induced by a signal. Its formal statement and its proof are in the appendix.

Lemma 2:

If the initial belief assigns a positive probability to at least three states of the world, then

(i) any signal with a positive intensity induces the most preferred decision to change.

(ii) the direction of this change depends only on the signal direction θ and not on the signal intensity a. It moves clockwise when θ does, which implies that there exists a one-to-one correspondence between signal directions and directions of the induced change of the most preferred decision.

(iii) as the signal intensity converges toward zero, the changes in the most preferred decision corresponding to the two possible signals tend to have the same magnitude.

Remark: Part (iii) reflects the fact that as *a* converges to zero, signals become uninformative and therefore equally likely *a priori*. The agent's expectation of the change of his most preferred decision being zero (it is a linear function of his expected change in belief), this implies that the two changes have the same magnitude and cancel each other in expectation.

The next lemma addresses a more difficult point: the orthogonality argument summarized in the introduction relies on the idea that a given signal changes different beliefs in similar ways, so that if agent 1, without any information, is indifferent between saying black or white, he should then prefer telling the truth when he knows it. This is not true in general. As the following example shows, agents starting with different beliefs may interpret the same signal in very different ways. If agent 1's belief assigns a positive probability only to the states of the world $(0,0)$ and $(0,1)$, and agent 2's only to $(0,0)$ and $(1,1)$, then observing a signal $S_{a,-\frac{\pi}{6}}$ moves agent's 1 most

preferred decision "downward", that is, the direction of the change of his most preferred decision is $-\frac{\pi}{2}$, while it moves agent 2's most preferred decision upward to the right

(the direction of this change is $\frac{\pi}{4}$), since for all $a > 0$, $f_{11}\left(S_{a,-\frac{\pi}{6}}\right) > f_{00}\left(S_{a,-\frac{\pi}{6}}\right)$.

Lemma 3 below qualifies this remark, and this will allow the orthogonal argument to work in the proof of Proposition 2. Its formal statement and its proof are in the appendix.

Lemma 3: *given a direction φ, there exists a signal direction moving the most preferred decision of one of the groups along the direction φ, and that of the other group along a direction forming an angle with φ smaller than $\frac{\pi}{2}$.*

Notice that Lemma 3 introduces a possible asymmetry between agents: given φ, it may be possible to move group 1 along the direction φ, and group 2 along a direction forming an acute angle with φ, while the converse may be false, as the above example shows.

We can now state the central result of the paper.

Proposition 2: *Assume (μ^1, μ^2) is such that μ^1 and μ^2 assign a strictly positive probability to more than three states of the world (that is, none of these beliefs belongs to an edge of Δ). Then (μ^1, μ^2) is not a steady-state.*

Sketch of the proof:

The exact proof is in the appendix. The argument summarized here is illustrated on Figure 1. Let us write M_i for $M^*(\mu^i)$ (the most preferred decision of group i agents before any communication takes place). Lemma 3 implies that, writing M_i^+ for $M^*(B(\mu^i, S_{a,\theta}))$ and M_i^- for $M^*(B(\mu^i, S_{a,\theta+\pi}))$, there exists a signal direction θ such that

- $M_2^- M_2^+$ (for example) is orthogonal to $M_1 M_2$
- M_1^+ and M_2^+ are located on the same side of the $M_1 M_2$.
- Lemma 2 implies in addition that the distances $M_i M_i^+$ and $M_i M_i^-$ are roughly equal if a is small.

13

One can check that this implies that

- $M_1^+ M_2^+ < M_1^+ M_2^-$ and
- $M_1^- M_2^- < M_1^- M_2^+$

Therefore a group 1 agent wants to truthfully report $S_{a,\theta}$ if he observed $S_{a,\theta}$ and expects to be believed: his most preferred decision is M_1^+ and the decision chosen by the receiver in group 2 is going to be M_2^+ if he reports $S_{a,\theta}$, M_2^- otherwise. The form of the indirect utility function stated in Lemma 1 implies that the sender wants to minimize the distance between his most preferred decision and the receiver's. The first inequality above implies that this is done by reporting $S_{a,\theta}$. Similarly, the second inequality implies that a group 1 agent wants to truthfully report $S_{a,\theta+\pi}$ if he observed $S_{a,\theta+\pi}$.

Therefore there exists a communicative equilibrium when the signal direction is θ, the intensity is small and the sender is in group 1. By continuity there exist many in a neighborhood around θ, so they occur with a strictly positive probability. This means that with probability 1, some communication will take place in some future period, and beliefs are going to change. Therefore (μ^1, μ^2) is not a steady-state. **Q.e.d.**
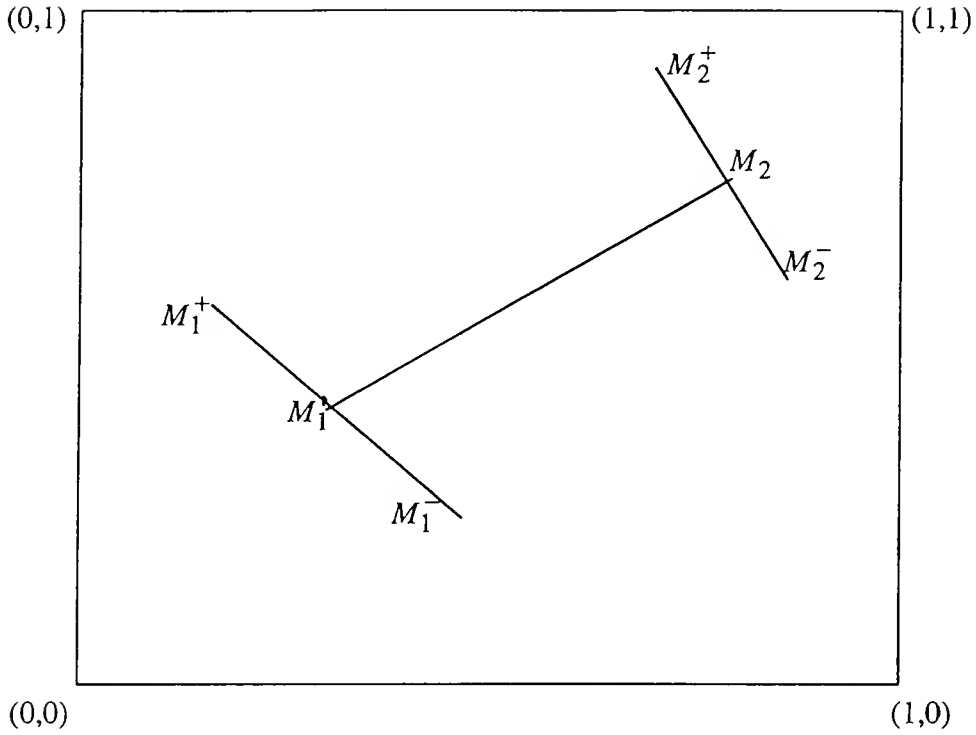
Figure 1

4. Dynamics, stability and convergence

For any initial condition (μ^{10}, μ^{20}), the stochastic process (μ^{1t}, μ^{2t}) is a bounded martingale, just as any bayesian learning process. That is, conditional on a member of group i's belief μ^{it} at time t, $E(\mu^{it+1}|\mu^{it}) = \mu^{it}$. Therefore the martingale convergence theorem[8] applies, and as t goes to infinity (μ^{1t}, μ^{2t}) converges toward

---

[8]See for example Neveu (1975).

15

some steady-state (μ^1, μ^2) with probability 1[9]. Having characterized the set of steady states in Section 3, we examine in this section how the probability distribution of limit beliefs on that set varies as a function of initial beliefs.

The description of the belief dynamics starts with the following restriction about possible limits:

Proposition 3: *Assume that* μ^{10} *and* μ^{20} *have full support. Then the limit belief distribution* (μ^1, μ^2) *is such that*

(i) *either* μ^1 *and* μ^2 *assign probability 1 to the true state of the world s, i.e.* $\mu^1 = \mu^2 = \delta_s$.

(ii) *or* μ^1 *and* μ^2 *both belong to some edge* $[\delta_s, \delta_w]$ *of* Δ *containing the true belief.*

Proof: A general result about bayesian learning is that if a belief assigns initially a strictly positive probability to the true state of the world, then with probability 1 its limit does as well[10]. If case (i) does not hold, there exists a state of the world $w \neq s$ such that $\mu_w^1 > 0$. Since almost all agents in both groups (all except those who directly observed a signal) received the same information between periods 0 and ∞, Bayes' rule implies that

$$\frac{\mu_w^1 \mu_s^{10}}{\mu_w^{10} \mu_s^1} = \frac{\mu_w^2 \mu_s^{20}}{\mu_w^{20} \mu_s^2} = \underset{t \to \infty}{Lim} \frac{\Pr(\text{messages up to } t|w)}{\Pr(\text{messages up to } t|s)}.$$

Therefore $\mu_w^2 > 0$. We know from Proposition 2 that neither μ^1 nor μ^2 can be in the interior of Δ, for (μ^1, μ^2) would not be a steady-state then. Therefore both μ^1 and μ^2 belong to $[\delta_s, \delta_w]$. **Q.e.d.**

---

[9]Strictly speaking, the martingale convergence theorem only implies that the system will converge somewhere with probability 1. The limit has to be a steady-state (in the sense defined in section 3 above) only if the transition correspondence has adequate continuity properties, which is the case if we select the most informative equilibrium at each stage (selecting the truth-telling equilibrium whenever it exists defines a continuous transition correspondence).

[10] Aghion et al. (1991).

16

The second case corresponds to convergence toward the steady states described in Proposition 1, characterized by a one-dimensional conflict. Although these are not the only steady states, Proposition 3 rules out convergence to any other (for example belief distributions such that beliefs belong to different edges). There remains to prove that one-dimensional conflict indeed happens at the limit with a strictly positive probability.

Although we do not fully describe the dynamics, we focus below on the following two cases: the one where the initial belief distribution is close to a one-dimensional steady-state (Proposition 4), and the one where initial beliefs are close to each other (Proposition 5).

Proposition 4 shows that if the initial belief distribution is close to a one-dimensional steady-state, then there is positive probability of converging toward a nearby one, and that this probability converges to 1 as the initial belief distribution converges to such a steady-state. This result is important: it means that the steady-states of Proposition 1 are locally stable (as long as they assign a positive probability to the truth): if such a steady-state is locally perturbed, the system converges back to a similar one with positive probability. This is a weak notion of stability, since we do not prove that this probability is equal to 1 for small perturbations – we only prove that it converges to 1 as the size of the perturbation converges toward zero. This proposition implies that the steady-states of Proposition 1 are meaningful: convergence toward a one-dimensional conflict occurs with a positive probability for a positive-measure set of initial beliefs.

Proposition 4: *Consider an edge L of Δ containing the true belief δ_s, a steady state $(\mu *^1, \mu *^2)$ in the interior of L^2, and a neighborhood W of $(\mu *^1, \mu *^2)$ in L^2.*

1. *There exists a neighborhood V of $(\mu *^1, \mu *^2)$ in Δ^2 such that if $(\mu^{10}, \mu^{20}) \in V$ then with a strictly positive probability $(\mu^{1t}, \mu^{2t})_{t \geq 0}$ converges toward a limit belonging to W.*
2. *This probability converges to 1 as (μ^{10}, μ^{20}) converge toward $(\mu *^1, \mu *^2)$.*

Sketch of the proof: The exact proof is in the Appendix. Its main lines are as follows (see figure 2): assume that steady state beliefs assign a zero probability to the event "$y_s = 1$" (the proof is less intuitive if steady state beliefs are correlated across dimensions). If $(\mu *^1, \mu *^2)$ is a steady state corresponding to most preferred decisions

17

$\left(M_1^*, M_2^*\right)$, then one can check that beliefs inducing most preferred decisions close to $\left(M_1^*, M_2^*\right)$ must be close to $(\mu*^1, \mu*^2)$. One shows first that such beliefs are characterized by "almost vertical" communication (Result 2 of the proof): horizontal communication is infeasible at $(\mu*^1, \mu*^2)$, and this implies by a continuity argument that as long as most preferred decisions are respectively in the interior of the squares $A_1 B_1 C_1 D_1$ and $A_2 B_2 C_2 D_2$, communication can cause them to move only along directions very close to the vertical. A standard property of martingales implies that if initial beliefs assign a very small probability to the event "$y_s = 1$", then with a large enough probability this probability remains small forever (Result 4). This means that if initially the most preferred decisions are in the small squares around M_1^* and M_2^*, then with a large enough probability, the most preferred decisions cannot leave the squares $A_1 B_1 C_1 D_1$ and $A_2 B_2 C_2 D_2$ by moving across the lines $B_1 C_1$ and $B_2 C_2$. But they cannot either cross the vertical lines $A_1 B_1$ or $A_2 B_2$: the fact that communication is almost along the vertical direction implies indeed that a given total horizontal movement corresponds to a total vertical "oscillation" of comparable magnitude, but the total vertical oscillation must remain very small (with a large probability) if the initial probability assigned to the event "$y_s = 1$" is small, and in particular if initially the most preferred decisions are in the small squares around M_1^* and M_2^* (this is shown in Result 4, building on the theorem about the quadratic variation of positive martingales). This implies that if initially the most preferred decisions are in these small squares, then with a positive probability, the most preferred decisions never leave the "large" squares $A_1 B_1 C_1 D_1$ and $A_2 B_2 C_2 D_2$. Belief convergence and Proposition 2 then imply that the limit distribution must be characterized by a one-dimensional disagreement, with group 1's belief lying between A_1 and D_1 and group 2's belief between A_2 and D_2.
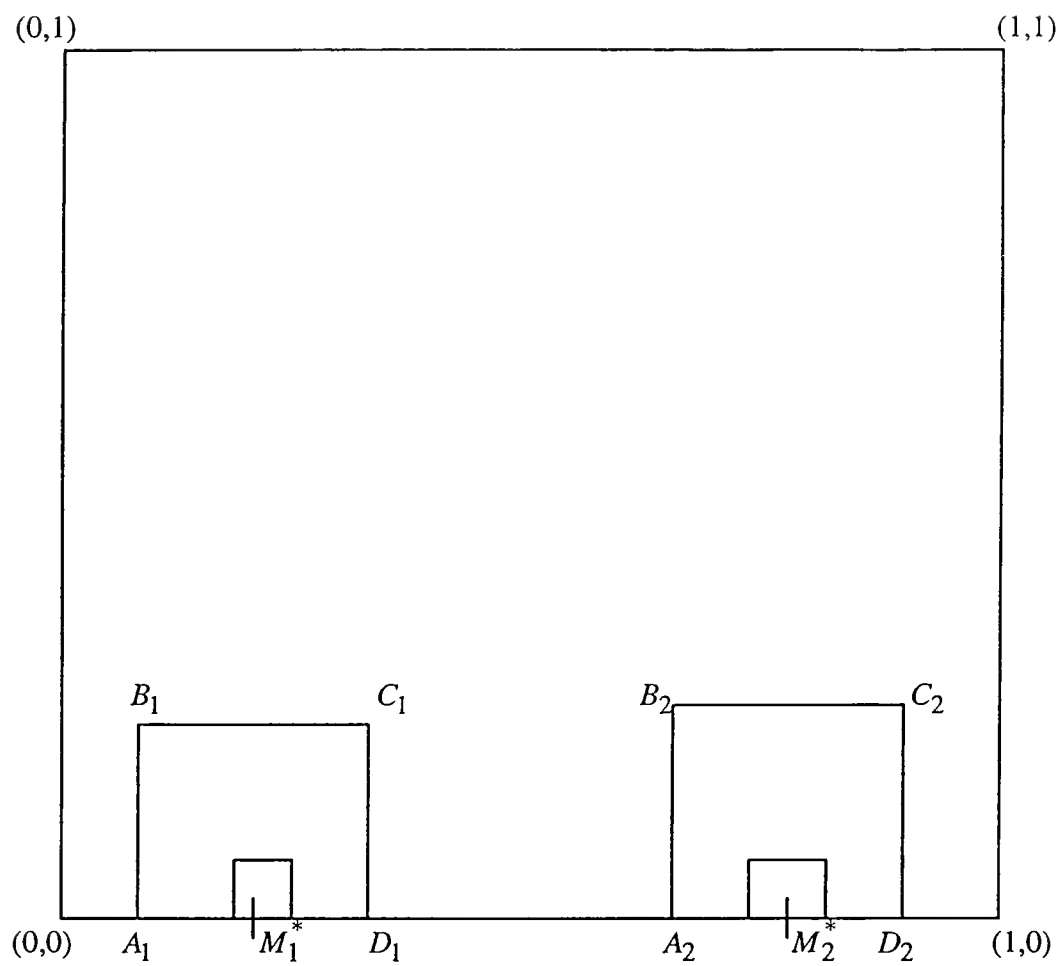
Figure 2

The next result qualifies Proposition 4: if initial beliefs are close enough to each other, then convergence toward the truth – and total agreement - occurs with probability 1.

Proposition 5: *If the ratios* $\dfrac{\mu_p^{10}\mu_q^{20}}{\mu_p^{20}\mu_q^{10}}$ *(for all pairs (p,q) of states of the world) are close enough to 1, then all beliefs converge towards the truth with probability one.*

Proof: Appendix

This result is stronger than a continuity property around identical initial beliefs: one could have expected the probability of convergence to the truth to converge to one, instead of being equal to one around a given initial distribution with identical beliefs. The proof relies on the following idea: since in period t almost all agents in both groups observe the same information (the sequence of messages up to period t), Bayes' rule implies that $\dfrac{\mu_p^{1t}\mu_q^{2t}}{\mu_p^{2t}\mu_q^{1t}}$ is independent of t, so that if it is close to 1 initially, it must also be close to 1 at any steady-state, meaning that the limit beliefs are close to each other. This is impossible, as one can show, with arguments similar to those of the proof of Proposition 1, that one-dimensional conflict requires some distance between beliefs.

To summarize, we have shown in this section that convergence toward one-dimensional conflict occurs with a positive probability for a positive measure-set (though not the entire set) of initial belief distributions.

5. Concluding remarks

One may wonder how sensitive the results are to the various assumptions. As explained in Section 2.4, they would fail for obvious reasons if there were a finite number of agents, or if an agent could speak to his own group only. We have a less clear idea about the possibility of relaxing the assumption of infinite impatience, as the forward-looking behavior this would induce makes the analysis of the communication game difficult (at the very least, a continuity argument should allow our results to carry over to the case of very impatient agents, with the set of possible steady-states converging toward the one found here as the rate of time preference goes to infinity).

20

Similarly, we do not know to which class of utility functions our results can be extended. The reason of this uncertainty is the very counterintuitive nature of the geometry of bayesian updating in several dimensions, exemplified by the fact that the changes of the most preferred decisions of two agents having observed the same signal can go in very different directions (Section 3). These problems may weaken the orthogonality argument. Extending the results to more than two groups would be particularly welcome: it would make the link with the evidence from Political Science more meaningful. The difficulty, here again, lies with generalizing lemma 3 - establishing that updating directions can be made close enough to each other. It should be possible, at least, to extend our results, by continuity, to the case where the initial belief distribution is close to two main groups - for example if all groups except two are small enough, or if all initial beliefs can be partitioned into two sets of neighboring beliefs.

Generalizing the model to more than two dimensions seems relatively straightforward: after extending the signal structure in a natural way, it should be sufficient to consider a plane containing both beliefs in the corresponding simplex, and to apply Proposition 2 to that plane by restricting the attention to signals leaving beliefs in that plane. This would show that steady-state beliefs must belong to a common edge of the simplex.

Further developments in bayesian theory might allow to assess how general our argument is, and how relevant it may be in terms of political economy. This should be the subject of future research.

Appendix

Proof of Lemma 1:

Given a belief μ about the state of the world, the expected utility induced by the decision M, is equal to (writing M*(μ) for $E\left(M_s|\mu\right)$)

$$V(M,\mu) = -\sum_s \mu_s \left(\overrightarrow{MM_s} \right)^2 = -\sum_s \mu_s \left(\overrightarrow{MM*}(\mu) + \overrightarrow{M*(\mu)M_s} \right)^2$$

$$= -\sum_s \mu_s \left(\left\| \overrightarrow{MM*}(\mu) \right\|^2 + 2\,\overrightarrow{MM*}(\mu).\overrightarrow{M*(\mu)M_s} + \left\| \overrightarrow{M*(\mu)M_s} \right\|^2 \right)$$

$$= -\left\| \overrightarrow{MM*}(\mu) \right\|^2 - \sum_s \mu_s \left\| \overrightarrow{M*(\mu)M_s} \right\|^2$$

which yields the result. Q.e.d.

Lemma 2:

If μ assigns a strictly positive probability to at least three states of the world then

(i) for all $a>0$ and θ, $M*(B(\mu,S_{a,\theta})) \neq M*(\mu)$. The direction of the vector $\overrightarrow{M*(\mu)M*(B(\mu,S_{a,\theta}))}$ is independent of a and moves clockwise as θ does. We denote it $\varphi_\mu(\theta)$ hereafter. φ_μ is a one-to-one function from $[0,2\pi]$ into itself.

(ii) $\dfrac{\partial}{\partial a}\Big|_{a=0} \left(M*(B(\mu,S_{a,\theta})) \right) \neq 0$.

(iii) $\dfrac{\partial}{\partial a}\Big|_{a=0} \left(\overrightarrow{M*(\mu)M*(B(\mu,S_{a,\theta}))} + \overrightarrow{M*(\mu)M*(B(\mu,S_{a,\theta+\pi}))} \right) = 0$.

Proof of Lemma 2:

We consider a belief $\mu = \left(\mu_{00},\mu_{10},\mu_{01},\mu_{11}\right)$ assigning a positive weight to at least three states of the world.

$$M*(\mu) = E\big(M_s\big|\mu\big) = (\mu_{10} + \mu_{11}, \mu_{01} + \mu_{11}).$$

The coordinates of M*(μ) are

$$\begin{cases} x(\mu) = \mu_{10} + \mu_{11} \\ y(\mu) = \mu_{01} + \mu_{11} \end{cases}$$

Bayes' rule implies that if an agent's initial belief was $\mu = \big(\mu_{00}, \mu_{10}, \mu_{01}, \mu_{11}\big)$ and he observed the signal $S_{a,\theta}$, then his new belief $\mu' = B(\mu, S_{a,\theta})$ is given by

$$\begin{cases} \Sigma\mu'_{00}(S_{a,\theta}) = \mu_{00}\big(1 + 2a(-\cos\theta - \sin\theta)\big) \\ \Sigma\mu'_{01}(S_{a,\theta}) = \mu_{01}\big(1 + 2a(-\cos\theta + \sin\theta)\big) \\ \Sigma\mu'_{10}(S_{a,\theta}) = \mu_{10}\big(1 + 2a(\cos\theta - \sin\theta)\big) \\ \Sigma\mu'_{11}(S_{a,\theta}) = \mu_{11}\big(1 + 2a(-\cos\theta - \sin\theta)\big) \end{cases} \qquad (1)$$

with Σ (a function of a and θ) is given by

$$\Sigma = \mu_{00}\big(1 + 2a(-\cos\theta - \sin\theta)\big) + \mu_{01}\big(1 + 2a(-\cos\theta + \sin\theta)\big)$$
$$+ \mu_{10}\big(1 + 2a(\cos\theta - \sin\theta)\big) + \mu_{11}\big(1 + 2a(-\cos\theta - \sin\theta)\big) \qquad (2)$$

Notice that the assumption $A < \dfrac{1}{2\sqrt{2}}$ ensures that $\Sigma > 0$.

The change in M*(μ) induced by observing $S_{a,\theta}$, denoted $\Delta M* = (\Delta x, \Delta y)$ is equal to

$$\big(\mu_{10}(S_{a,\theta}) + \mu_{11}(S_{a,\theta}) - \mu_{10} - \mu_{11}, \mu_{01}(S_{a,\theta}) + \mu_{11}(S_{a,\theta}) - \mu_{01} - \mu_{11}\big).$$

Substituting (1) and (2) in this expression yields

$$\begin{cases} \dfrac{\Sigma(\Delta x)}{a} = & (\cos\theta + \sin\theta)\big[(\mu_{11} + \mu_{00}) - (\mu_{11} - \mu_{00})(\mu_{11} - \mu_{00} + \mu_{01} - \mu_{10})\big] \\ & + (\cos\theta - \sin\theta)\big[(\mu_{10} + \mu_{01}) - (\mu_{10} - \mu_{01})(\mu_{10} - \mu_{01} - \mu_{00} + \mu_{11})\big] \\ \dfrac{\Sigma(\Delta y)}{a} = & (\cos\theta + \sin\theta)\big[(\mu_{11} + \mu_{00}) - (\mu_{11} - \mu_{00})(\mu_{11} - \mu_{00} - \mu_{01} + \mu_{10})\big] \\ & - (\cos\theta - \sin\theta)\big[(\mu_{10} + \mu_{01}) - (\mu_{10} - \mu_{01})(\mu_{10} - \mu_{01} + \mu_{00} - \mu_{11})\big] \end{cases} \qquad (3)$$

<u>Result 1</u>: *We assume that* $\theta \in \left[0, \dfrac{\pi}{4}\right]$. *Then* $\Delta x > 0$, $\Delta x + \Delta y > 0$. *This implies that* $(\Delta x, \Delta y) \neq (0,0)$ *for any* $\theta \in \left[0, \dfrac{\pi}{4}\right]$. *In addition the direction of* $(\Delta x, \Delta y)$ *is independent of* a.

<u>Proof:</u>

- For $\theta \in \left[0, \dfrac{\pi}{4}\right]$, (3) implies indeed that

$$
\frac{\Sigma.(\Delta x + \Delta y)}{a(\cos\theta + \sin\theta)} = \left[(\mu_{11} + \mu_{00}) - (\mu_{11} - \mu_{00})^2\right] - \frac{(\cos\theta - \sin\theta)}{\cos\theta + \sin\theta}(\mu_{11} - \mu_{00})(\mu_{10} - \mu_{01})
$$

$$
\geq \left[(\mu_{11} + \mu_{00}) - (\mu_{11} + \mu_{00})^2\right] - \frac{(\cos\theta - \sin\theta)}{\cos\theta + \sin\theta}(\mu_{11} - \mu_{00})(\mu_{10} - \mu_{01})
$$

$$
= (\mu_{11} + \mu_{00})(\mu_{10} + \mu_{01}) - \frac{(\cos\theta - \sin\theta)}{\cos\theta + \sin\theta}(\mu_{11} - \mu_{00})(\mu_{10} - \mu_{01}).
$$

If three at least of the probabilities $\mu_{00}, \mu_{10}, \mu_{01}, \mu_{11}$ are strictly positive, then $(\mu_{11} + \mu_{00})(\mu_{10} + \mu_{01}) > (\mu_{11} - \mu_{00})(\mu_{10} - \mu_{01})$. Also, $1 \geq \dfrac{(\cos\theta - \sin\theta)}{\cos\theta + \sin\theta}$ if $\theta \in \left[0, \dfrac{\pi}{4}\right]$.
Therefore

$$
(\mu_{11} + \mu_{00})(\mu_{10} + \mu_{01}) > \frac{(\cos\theta - \sin\theta)}{\cos\theta + \sin\theta}(\mu_{11} - \mu_{00})(\mu_{10} - \mu_{01})
$$

and $\Delta x + \Delta y > 0$.

- Similarly, for $\theta \in \left[0, \dfrac{\pi}{4}\right]$, (3) implies that

$$
\frac{\Sigma.(\Delta x)}{a} = (\cos\theta + \sin\theta)\left[(\mu_{11} + \mu_{00}) - (\mu_{11} - \mu_{00})(\mu_{11} - \mu_{00} - \mu_{10} + \mu_{01}))\right]
$$

$$
+ (\cos\theta - \sin\theta)\left[(\mu_{10} + \mu_{01}) - (\mu_{10} - \mu_{01})(\mu_{10} - \mu_{01} + \mu_{11} - \mu_{00}))\right].
$$

If three at least of the probabilities $\mu_{00}, \mu_{10}, \mu_{01}, \mu_{11}$ are strictly positive, then $|\mu_{11} - \mu_{00} - \mu_{10} + \mu_{01}| < 1$ and $|\mu_{10} - \mu_{01} + \mu_{11} - \mu_{00}| < 1$ so that the factors in brackets

are strictly positive. The inequalities $\cos\theta + \sin\theta > 0$ and $\cos\theta - \sin\theta \geq 0$ imply then that $\Delta x > 0$.

- Any of these two strict inequalities implies that $(\Delta x, \Delta y) \neq (0,0)$. In addition (3) implies that the direction of $(\Delta x, \Delta y)$ is independent of a. **Q.e.d.**

Result 2: *For every θ in $[0,2\pi]$ the change $(\Delta x, \Delta y)$ of $M^*(\mu)$ induced by observing $S_{a,\theta}$ is different from $(0,0)$ and its direction does not depend on a. If one writes $\varphi_\mu(\theta)$ for this direction, then for every integer k*

$$(i) \quad \theta \in \left[\frac{k\pi}{4}, \frac{(k+1)\pi}{4} \right] \text{ implies that } \varphi_\mu(\theta) \in \left(\frac{(k-1)\pi}{4}, \frac{(k+2)\pi}{4} \right)$$

and

$$(ii) \quad \varphi_\mu\left(\frac{k\pi}{4} \right) \in \left(\frac{(k-1)\pi}{4}, \frac{(k+1)\pi}{4} \right).$$

Proof:

By Result 1 we know that if $\theta \in \left[0, \frac{\pi}{4} \right]$ then $\Delta(x+y) > 0$ and $\Delta x > 0$, implying respectively $\varphi_\mu(\theta) \in \left(-\frac{\pi}{4}, \frac{5\pi}{4} \right)$ and $\varphi_\mu(\theta) \in \left(-\frac{\pi}{2}, \frac{\pi}{2} \right)$. These two inclusions imply that if $\theta \in \left[0, \frac{\pi}{4} \right]$ then $\varphi_\mu(\theta) \in \left(-\frac{\pi}{4}, \frac{\pi}{2} \right)$. We proved (i) for k=0, and it remains true for all integers k because of the symmetry of the model.

In particular $\varphi_\mu(0) > -\frac{\pi}{4}$ and $\varphi_\mu\left(\frac{\pi}{4} \right) < \frac{\pi}{2}$.

By symmetry these two inequalities have the following counterparts:

$\varphi_\mu(0) < \frac{\pi}{4}$ and $\varphi_\mu\left(\frac{\pi}{4} \right) > 0$, so that $\varphi_\mu\left(\frac{k\pi}{4} \right) \in \left(\frac{(k-1)\pi}{4}, \frac{(k+1)\pi}{4} \right)$ is true for k=0 and k=1.

25

Again, the symmetry argument implies that it is true for every integer k. **Q.e.d.**

Result 3: *We assume that* $\theta \in \left[0, \dfrac{\pi}{4}\right]$. *Then* $\dfrac{\Delta y}{\Delta x}$ *is a increasing function of* θ.

Proof:

Notice first that by Result 1 $\dfrac{\Delta y}{\Delta x}$ is always defined if $\theta \in \left[0, \dfrac{\pi}{4}\right]$. (3) and the fact that the derivative of the function $\dfrac{B(\cos\theta + \sin\theta) - C(\cos\theta - \sin\theta)}{D(\cos\theta + \sin\theta) + E(\cos\theta - \sin\theta)}$ has the same sign as BE+CD implies that the derivative of $\dfrac{\Delta y}{\Delta x}$ has the same sign as

$$[(\mu_{11} + \mu_{00}) - (\mu_{11} - \mu_{00})(\mu_{11} - \mu_{00} + \mu_{01} - \mu_{10})][(\mu_{10} + \mu_{01}) - (\mu_{10} - \mu_{01})(\mu_{10} - \mu_{01} + \mu_{00} - \mu_{11})]$$
$$+ [(\mu_{10} + \mu_{01}) - (\mu_{10} - \mu_{01})(\mu_{10} - \mu_{01} - \mu_{00} + \mu_{11})][(\mu_{11} + \mu_{00}) - (\mu_{11} - \mu_{00})(\mu_{11} - \mu_{00} - \mu_{01} + \mu_{10})]$$

But the term in each bracket is strictly positive if three at least of the μ_{ij} are strictly positive, so the whole expression is positive and $\dfrac{\Delta y}{\Delta x}$ is a increasing function of θ. **Q.e.d.**

Result 4: *The function* $\varphi_\mu(\theta)$ *is a one-to-one correspondence from [0,2π] into itself, and* $\varphi_\mu(\theta)$ *moves clockwise as θ does.*

Proof: If $\theta \in \left[0, \dfrac{\pi}{4}\right]$, $\varphi_\mu(\theta)$ belongs to $\left(-\dfrac{\pi}{4}, \dfrac{\pi}{2}\right)$ by Result 2. Since $\varphi_\mu(\theta)$ is the direction of the vector $(\Delta x, \Delta y)$, $\dfrac{\Delta y}{\Delta x} = \tan(\varphi_\mu(\theta))$. Result 4 implies then that $\varphi_\mu(\theta)$ is increasing with θ, and that $\varphi_\mu(\theta)$ moves clockwise as θ does. By symmetry, this is true for $\theta \in [0, 2\pi]$.

This implies that the function $\varphi_\mu(\theta)$ is a one-to-one correspondence, since it comes back to the same value as θ moves from 0 to 2π. **Q.e.d.**

Result 5: *For any θ,* $\dfrac{\partial}{\partial a}\left(M*(B(\mu, S_{a,\theta}))\right)\bigg|_{a=0} \neq 0$.

Proof:

Again, consider $\theta \in \left(0, \dfrac{\pi}{4}\right]$. (3) implies

$$\frac{\Sigma(a,\theta)(\Delta x)}{a} = \begin{aligned}[t] &(\cos\theta + \sin\theta)\left[(\mu_{11} + \mu_{00}) - (\mu_{11} - \mu_{00})(\mu_{11} - \mu_{00} + \mu_{01} - \mu_{10})\right] \\ &+ (\cos\theta - \sin\theta)\left[(\mu_{10} + \mu_{01}) - (\mu_{10} - \mu_{01})(\mu_{10} - \mu_{01} - \mu_{00} + \mu_{11})\right] \end{aligned}$$

The right-hand side of this equality is independent of a and different from zero (as proved in Result 1), let us denote it $C(\theta)$.

Differentiating the above equality at $a=0$ yields

$$\left.\frac{\partial}{\partial a}(\Delta x)\right|_{a=0} = \frac{C(\theta)}{\Sigma(0,\theta)} = C(\theta) \neq 0. \textbf{ Q.e.d.}$$

Result 6: *For any θ,* $\left.\dfrac{\partial}{\partial a}\left(\Delta M * (B(\mu, S_{a,\theta})) + \Delta M * (B(\mu, S_{a,\theta+\pi}))\right)\right|_{a=0} = (0,0)$.

Proof:

We proved above that $\left.\dfrac{\partial}{\partial a}\left(\Delta x(S_{a,\theta})\right)\right|_{a=0} = C(\theta)$.

Therefore $\left.\dfrac{\partial}{\partial a}\left(\Delta x(S_{a,\theta}) + \Delta x(S_{a,\theta+\pi})\right)\right|_{a=0} = C(\theta) + C(\theta + \pi) = 0$.

Similarly $\left.\dfrac{\partial}{\partial a}\left(\Delta y(S_{a,\theta}) + \Delta y(S_{a,\theta+\pi})\right)\right|_{a=0} = 0$, implying

$$\left.\frac{\partial}{\partial a}\left(\Delta M * (B(\mu, S_{a,\theta})) + \Delta M * (B(\mu, S_{a,\theta+\pi}))\right)\right|_{a=0}$$

$$= \left(\left.\frac{\partial}{\partial a}\left(\Delta x(S_{a,\theta}) + \Delta x(S_{a,\theta+\pi})\right)\right|_{a=0}, \left.\frac{\partial}{\partial a}\left(\Delta y(S_{a,\theta}) + \Delta y(S_{a,\theta+\pi})\right)\right|_{a=0}\right) = (0,0). \textbf{ Q.e.d.}$$

Results 1 to 6 prove Lemma 2: part (i) is just Results 2 and 4, (ii) is Result 5, and (iii) is Result 6.

Lemma 3: *If μ_1 and μ_2 assign a strictly positive probability to at least three states of the world, then for any direction φ in $[0,2\pi]$, we have (writing $\theta_\mu(\varphi)$ for the inverse function of φ_μ):*

$$Min\left(\left|\varphi_{\mu_1}(\theta_{\mu_2}(\varphi)) - \varphi\right|, \left|\varphi_{\mu_2}(\theta_{\mu_1}(\varphi)) - \varphi\right|\right) < \frac{\pi}{2}.$$

In other words, one at least of the scalar products

$$M*(\mu_1)M*(\overrightarrow{B(\mu_1,S_{\theta_i(\varphi)})}).\, M*(\mu_2)M*(\overrightarrow{B(\mu_2,S_{\theta_i(\varphi)})})$$

is strictly positive (i=1 or i=2).

Proof of Lemma 3:

We use the same notations as in the proof of Lemma 2.

Result 1: *If $\theta = \dfrac{k\pi}{2}$ (k integer) and $(\Delta x, \Delta y)$ is the change of M*(μ) induced by the observation of $S_{a,\theta}$, then the product $\Delta x.\Delta y$ has the same sign as the correlation of μ across dimensions, that is, the same sign as $\mu_{11}\mu_{00} - \mu_{01}\mu_{10}$. This implies that if $\mu_{11}\mu_{00} - \mu_{01}\mu_{10}$ is positive (resp. negative), then $\varphi_\mu(0) - 0$ and $\varphi_\mu(\pi) - \pi$ are positive (negative) while $\varphi_\mu\left(\dfrac{\pi}{2}\right) - \dfrac{\pi}{2}$ and $\varphi_\mu\left(\dfrac{3\pi}{2}\right) - \dfrac{3\pi}{2}$ are negative.*

Proof: It is enough to prove the result for k=0. We know already (from Result 2 in the proof of Lemma 2) that

$\varphi_\mu(0) \in \left(-\dfrac{\pi}{4}, \dfrac{\pi}{4}\right)$. Since Δy and $\varphi_\mu(0) - 0$ have the same sign, the only thing we need to prove is that if $\theta=0$, then Δy has the same sign as $\mu_{11}\mu_{00} - \mu_{01}\mu_{10}$.

28

Substituting θ=0 in (3) (in the proof of Lemma 2) implies that

$$\Delta y = \mu_{11} + \mu_{00} - (\mu_{11} - \mu_{00})^2 - (\mu_{10} + \mu_{01}) + (\mu_{10} - \mu_{01})^2$$
$$= (\mu_{11} + \mu_{00}) - (\mu_{11} + \mu_{00})^2 - (\mu_{10} + \mu_{01}) + (\mu_{10} + \mu_{01})^2 + 4(\mu_{11}\mu_{00} - \mu_{10}\mu_{01})$$
$$= (\mu_{11} + \mu_{00})(\mu_{10} + \mu_{01}) - (\mu_{11} + \mu_{00})(\mu_{10} + \mu_{01}) + 4(\mu_{11}\mu_{00} - \mu_{10}\mu_{01})$$
$$= 4(\mu_{11}\mu_{00} - \mu_{10}\mu_{01})$$

(the third equality uses the fact that the probabilities add up to 1),

which yields the result. **Q.e.d.**

This result is intuitive: if a signal providing information about only one dimension changes the belief regarding the other dimension, then the changes along the two dimensions must have the same (opposite) sign if the initial belief is characterized by a positive (negative) correlation across dimensions.

Result 2: *Lemma 3 holds with* $\varphi \in \left[0, \dfrac{\pi}{4}\right]$.

We know from Result 2 (in the proof of Lemma 2) that $\theta_1(\varphi)$ and $\theta_2(\varphi)$ belong to $\left(-\dfrac{\pi}{4}, \dfrac{\pi}{2}\right)$.

We distinguish three cases, according to the correlation of μ_1 and μ_2 across dimensions.

- If μ_1 and μ_2 display positive correlation across dimensions, then by Result 1 $\varphi_i\left(\dfrac{\pi}{2}\right) \le \dfrac{\pi}{2}$ (i=1,2) so that if for example $\theta_1(\varphi) \le \theta_2(\varphi)$ then by Result 4 (in the proof of Lemma 2) $0 \le \varphi = \varphi_1(\theta_1(\varphi)) \le \varphi_1(\theta_2(\varphi)) < \dfrac{\pi}{2}$ so that $\left|\varphi_1(\theta_2(\varphi)) - \varphi\right| < \dfrac{\pi}{2}$.

- If μ_1 and μ_2 display negative correlation across dimensions, then by Result 1 $\varphi_i(0) < 0$ (i=1,2) so that $\theta_i(\varphi) \ge \theta_i(0) \ge 0$ (i=1,2). If for example $\theta_1(\varphi) \le \theta_2(\varphi)$ then $\dfrac{\pi}{4} \ge \varphi = \varphi_2(\theta_2(\varphi)) \ge \varphi_2(\theta_1(\varphi)) \ge \varphi_2(0) > -\dfrac{\pi}{4}$ so that $\left|\varphi_2(\theta_1(\varphi)) - \varphi\right| < \dfrac{\pi}{2}$.

- If μ_1 and μ_2 display respectively positive and negative correlation then, by the same argument, $\dfrac{\pi}{2} > \theta_2(\varphi) \ge \theta_2(0) \ge 0$ and $\left[\varphi_1(0), \varphi_1\left(\dfrac{\pi}{2}\right)\right] \subset \left[0, \dfrac{\pi}{2}\right]$ (Result 1) implying that $0 \le \varphi_1\big(\theta_2(\varphi)\big) < \dfrac{\pi}{2}$ and $\big|\varphi_1(\theta_2(\varphi)) - \varphi\big| < \dfrac{\pi}{2}$. **Q.e.d.**

By symmetry, Lemma 3 holds for all values of φ. **Q.e.d.**

Proof of Proposition 2:

Let $\varphi \in [0, \pi)$ be the direction orthogonal to the vector $\overrightarrow{M^*(\mu^1)M^*(\mu^2)}$ if $M^*(\mu^1) \ne M^*(\mu^2)$, and any direction otherwise. Let us define for every θ the vector $\overrightarrow{V}(\theta) = (\cos\theta, \sin\theta)$. Using the notations of lemmas 2 and 3, we can assume for example that, with $\theta_2 = \theta_2(\varphi)$, the inequality

$$\overrightarrow{V(\varphi_1(\theta_2))}.\overrightarrow{V(\varphi_2(\theta_2))} = \overrightarrow{V(\varphi_1(\theta_2))}.\overrightarrow{V(\varphi)} > 0 \quad (1)$$

holds (this is an application of Lemma 3).

Parts (i) and (ii) of Lemma 2 implies the existence of λ,λ'>0 such that for a small enough

$$\begin{cases} M^*(B(\mu^2, S_{a,\theta_2+\pi}))\,\overrightarrow{M^*(\mu^2)} = a(\lambda + O(a))\overrightarrow{V(\varphi)} \\ M^*(\mu^2)\,\overrightarrow{M^*(B(\mu^2, S_{a,\theta_2}))} = a(\lambda + O(a))\overrightarrow{V(\varphi)} \end{cases}$$

and

$$\begin{cases} M^*(B(\mu^1, S_{a,\theta_2+\pi}))\,\overrightarrow{M^*(\mu^1)} = a(\lambda' + O(a))\overrightarrow{V(\varphi_1(\theta_2))} \\ M^*(\mu^1)\,\overrightarrow{M^*(B(\mu^1, S_{a,\theta_2}))} = a(\lambda' + O(a))\overrightarrow{V(\varphi_1(\theta_2))}. \end{cases}$$

30

We claim now that if the speaker belongs to group 1 and it is commonly known that the signal is S_{a,θ_2} or $S_{a,\theta_2+\pi}$ with a small enough, there exists a communicative equilibrium. By symmetry, it suffices to show that truth-telling is optimal for the speaker after he observed S_{a,θ_2}.

Writing M_i^+ for $M*(B(\mu^i, S_{a,\theta_2}))$, M_i^- for $M*(B(\mu^i, S_{a,\theta_2+\pi}))$ and M_i for $M*(\mu^i)$, the difference between the speaker's expected utility when he tells the truth (and is believed) and when he lied is

$$\left\| M_1^+ M_2^- \right\|^2 - \left\| M_1^+ M_2^+ \right\|^2 \qquad (2).$$

The identity $\left(\overrightarrow{v_1}\right)^2 - \left(\overrightarrow{v_2}\right)^2 = \left(\overrightarrow{v_1} - \overrightarrow{v_2}\right).\left(\overrightarrow{v_1} + \overrightarrow{v_2}\right)$ implies that

$$\left\| M_1^+ M_2^- \right\|^2 - \left\| M_1^+ M_2^+ \right\|^2 = \overrightarrow{M_2^+ M_2^-}.\left(\overrightarrow{M_1^+ M_2^+} + \overrightarrow{M_1^+ M_2^-}\right)$$

$$-2a(\lambda + O(a))\overrightarrow{V(\varphi)}.\left(2\overrightarrow{M_1 M_2} + 2\overrightarrow{M_1^+ M_1} + \overrightarrow{M_2 M_2^+} + \overrightarrow{M_2 M_2^-}\right)$$

$$-2a(\lambda + O(a))\overrightarrow{V(\varphi)}.\left(2\overrightarrow{M_1 M_2} - 2a(\lambda' + O(a))V(\overrightarrow{\varphi_1(\theta_2)}) + aO(a)\right)$$

$$= 4\lambda\lambda' a^2 \left(\overrightarrow{V(\varphi)}.V(\overrightarrow{\varphi_1(\theta_2)}) + O(a)\right)$$

which is strictly positive by (1) if a is small enough (the last equality results from the definition of $\overrightarrow{V(\varphi)}$, implying $\overrightarrow{M_1 M_2}.\overrightarrow{V(\varphi)} = 0$).

The expression in (2) and the corresponding expression in the case where the speaker observed $S_{a,\theta_2+\pi}$ are continuous functions of θ and a so they are strictly positive for an open (and therefore positive probability) subset of $(0,A]x[0,\pi)$. In the beginning of any odd period the speaker is in group 1 and there is a strictly positive probability that the signal's intensity and direction are going to be in this subset, so there is a strictly positive

probability that an informative equilibrium will occur. Therefore $\left(\mu^1, \mu^2\right)$ is not a steady-state. **Q.e.d.**

<u>Proof of Proposition 4:</u>

Let L be an edge of Δ containing the true belief δ_s. We consider a steady state $(\mu *^1, \mu *^2)$ in the interior of L^2. Note first that there exists a unique direction θ_L such that a signal along the direction θ_L does not affect beliefs belonging to L (it is the direction orthogonal to L). We write hereafter w for the state of the world such that $L = [\delta_s, \delta_w]$, and p,q for the two other states of the world.

Notations:

It will be convenient to denote the coordinates of a most preferred policy (an element of the square Q) with respect to the axis formed by [s,w] and [p,q]. These axis are either sides of Q (in case s and w agree on one dimension and disagree on the other) or a diagonal of Q. One can check that the coordinates u(μ) and v(μ) defined by

$$\begin{cases} u(\mu) = \dfrac{\mu_s - \mu_w}{2} \\ v(\mu) = \mu_p + \mu_q \end{cases}$$

if [s,w] and [p,q] are sides of Q, and by

$$\begin{cases} u(\mu) = 2^{-\frac{1}{2}} \left(\mu_s - \mu_w\right) \\ v(\mu) = 2^{-\frac{1}{2}} \left(\mu_p - \mu_q\right) \end{cases}$$

if [s,w] and [p,q] are diagonals of Q

are coordinates (modulo a constant) of $M*(\mu)$ according to these new axis. We will use this system of coordinates to describe elements of Q hereafter, so that one can write for example $M*(\mu) = (u(\mu), v(\mu))$ where u(μ) and v(μ) are defined as above.

Result 1: *As (μ^1, μ^2) converges towards $(\mu*^1, \mu*^2)$, the change in beliefs induced by communication converges toward 0.*

Proof: This is a simple continuity result. If this were not true, there would exist a sequence (μ^{1n}, μ^{2n}) converging towards $(\mu*^1, \mu*^2)$, a sequence of directions θ_n converging to some $\theta' \neq \theta_L, \theta_L + \pi$, and a sequence a_n converging toward some $a>0$ such that communication is possible with a speaker in group 1, or equivalently

$$\left\| M*(B(\mu^{1n}, S_{a_n,\theta})) M*(B(\mu^{2n}, S_{a_n,\theta})) \right\|$$
$$\leq \left\| M*(B(\mu^{1n}, S_{a_n,\theta})) M*(B(\mu^{2n}, S_{a_n,\theta+\pi})) \right\|$$

for $\theta = \theta_n$ and $\theta = \theta_n + \pi$.

This implies at the limit that

$$\left\| M^*(B(\mu*^1, S_{a,\theta})) M*(B(\mu*^2, S_{a,\theta})) \right\| \leq \left\| M^*(B(\mu*^1, S_{a,\theta})) M*(B(\mu*^2, S_{a,\theta+\pi})) \right\|$$

for $\theta = \theta'$ and $\theta = \theta'+\pi$ so that there exists a communicative equilibrium when the belief distribution is $(\mu*^1, \mu*^2)$, which is impossible since $(\mu*^1, \mu*^2)$ is a steady-state.

Result 2: *Consider the edge L'=[s,w] of [0,1]x[0,1] containing the most preferred policies associated to beliefs in L. Assume that $u(\mu*^2) - u(\mu*^1) = B > 0$. There exist a neighborhood V_0 of $(\mu*^1, \mu*^2)$ and some K>0 such that if $(\mu^1, \mu^2) \in V_0$, then for any equilibrium and realization of uncertainty, denoting the change of a group i agent's most preferred decision by $(\Delta u_i, \Delta v_i)$, the inequality*

$$Max[\Delta u_1, \Delta u_2] < K(Max[\Delta v_1, \Delta v_2])^2$$

holds.

Remark: this result implies that the direction of the changes of the most preferred decisions tend to become orthogonal to L. It is in fact stronger than that (orthogonality would only require that the ratio $\dfrac{Max[\Delta u_1, \Delta u_2]}{Max[\Delta v_1, \Delta v_2]}$ converge to zero).

Proof:

Assume for example that there an equilibrium where the speaker is in group 1, the signal intensity is a and the signal direction is θ. We denote $(\Delta x_i^+, \Delta y_i^+)$ (resp. $(\Delta x_i^-, \Delta y_i^-)$) the change in a group i agent's most preferred decision induced by the signal $S_{a,\theta}$ (resp. $S_{a,\theta+\pi}$). Since the most preferred decision is a linear function of the beliefs (it is the expectation) and bayesian updating implies that the expectation of next period's belief is the current belief, the expected change in the most preferred decision is zero, or

$$\Pr\!\left(S_{a,\theta}\middle|\mu^i\right)\!\left(\Delta u_i^+, \Delta v_i^+\right) + \Pr\!\left(S_{a,\theta+\pi}\middle|\mu^i\right)\!\left(\Delta u_i^-, \Delta v_i^-\right) = 0.$$

The signal structure is such that the ratios $\dfrac{\Pr\!\left(S_{a,\theta}\middle|\mu^i\right)}{\Pr\!\left(S_{a,\theta+\pi}\middle|\mu^i\right)}$ are bounded away from zero and infinity uniformly (that is, independently of a, θ and μ^i) as they belong to the interval $\left[\dfrac{1-2A\sqrt{2}}{1+2A\sqrt{2}}, \dfrac{1+2A\sqrt{2}}{1-2A\sqrt{2}}\right]$.

Therefore there exists some uniform K_1 such that

$$\frac{1}{K_1}\Delta u_i^- < \Delta u_i^+ < K_1\Delta u_i^- \text{ and } \frac{1}{K_1}\Delta v_i^- < \Delta v_i^+ < K_1\Delta v_i^-.$$

Let us suppose without loss of generality that $\Delta u_2^+ > 0$. The existence of a communicative equilibrium implies that a group 1 agent who observed $S_{a,\theta}$ (and whose most preferred decision moved by $(\Delta u_i^+, \Delta v_i^+)$) is better off by truthfully revealing he observed $S_{a,\theta}$ than by falsely reporting $S_{a,\theta+\pi}$, or

$$\left\|M^*(B(\mu^{*1}, S_{a,\theta}))M^*(B(\mu^{*2}, S_{a,\theta+\pi}))\right\|^2 - \left\|M^*(B(\mu^{*1}, S_{a,\theta}))M^*(B(\mu^{*2}, S_{a,\theta}))\right\|^2 \geq 0.$$

The left-hand side is equal to

$$\left\| \left(B + \Delta u_2^- - \Delta u_1^+, \Delta v_2^- - \Delta v_1^+\right) \right\|^2 - \left\| \left(B + \Delta u_2^+ - \Delta u_1^+, \Delta v_2^+ - \Delta v_1^+\right) \right\|^2$$

$$= \left(B + \Delta u_2^- - \Delta u_1^+\right)^2 + \left(\Delta v_2^- - \Delta v_1^+\right)^2 - \left(B + \Delta u_2^+ - \Delta u_1^+\right)^2 + \left(\Delta v_2^+ - \Delta v_1^+\right)^2$$

$$= \left(2B + \Delta u_2^- + \Delta u_2^+ - 2\Delta u_1^+\right)\left(\Delta u_2^- - \Delta u_2^+\right) + \left(\Delta v_2^- - \Delta v_2^+\right)\left(\Delta v_2^- + \Delta v_2^+ - \Delta v_1^- - \Delta v_1^+\right)$$

If (μ^1, μ^2) is close enough to $(\mu*^1, \mu*^2)$, Result 1 (and the facts that $\Delta u_2^+ > 0$ and Δu_2^+, Δu_2^- have opposite signs) implies that the first term is negative and greater in absolute value than $B\Delta u_2^+$. Similarly, the second term is smaller in absolute value than

$$2(1 + K_1)v_2^+ Max\left(\Delta v_2^+, \Delta v_1^+\right) \le 2(1 + K_1)\left[Max\left(\Delta v_2^+, \Delta v_1^+\right)\right]^2 .$$

The equilibrium condition implies that the second term is greater in absolute value than the first (since the first in negative and the condition states that their sum is positive), or

$$B\Delta u_2^+ \le 2(1 + K_1)\left[Max\left(\Delta v_2^+, \Delta v_1^+\right)\right]^2 .$$

Since $\mu*^1, \mu*^2$ are interior to $L = [\delta_s, \delta_w]$, there exists K_2 such that if (μ^1, μ^2) is close enough to $(\mu*^1, \mu*^2)$ then $\dfrac{1}{K_2}\Delta u_2^+ < \Delta u_1^+ < K_2 \Delta u_2^+$, so that the inequality above implies

$$BMax\left(\Delta u_1^+, \Delta u_2^+\right) \le 2(1 + K_1)(1 + K_2)\left[Max\left(\Delta v_2^+, \Delta v_1^+\right)\right]^2 .$$

This implies in turn

$$BMax\left(\Delta u_1^-, \Delta u_2^-\right) \le 2(K_1)^3(1 + K_1)(1 + K_2)\left[Max\left(\Delta v_2^-, \Delta v_1^-\right)\right]^2$$

so that the result holds with $K = 2(K_1)^3(1 + K_1)(1 + K_2)$.

Result 3: *Let us fix some $\eta, \varepsilon > 0$. There exists a neighborhood $V_1(\eta, \varepsilon)$ of $(\mu*^1, \mu*^2)$ such that if (μ^{10}, μ^{20}) belongs to $V_1(\eta, \varepsilon)$ then, writing $\left(\Delta u^{it}, \Delta v^{it}\right)$ for*

the change in the most preferred decision of a group i agent between periods t and t+1, the following inequality holds:

$$\Pr\left(\sum_{t \geq 0}\left(\Delta y^{it}\right)^2 > \eta\right) < \varepsilon.$$

Proof: Note first that if L consists of beliefs agreeing along one dimension (so that the corresponding set of most preferred decisions is a side of Q) then $\left|\Delta v^{it}\right| = \left|\Delta\mu_p^{it} + \Delta\mu_q^{it}\right|$ while if [p,q] is a diagonal then then $\left|\Delta v^{it}\right| = 2^{-\frac{1}{2}}\left|\Delta\mu_p^{it} - \Delta\mu_q^{it}\right|$. In both cases, the following inequality holds:

$$\sum_{t \geq 1}\left(\Delta v^{it}\right)^2 < 2\sum_{t \geq 1}\left(\left(\Delta\mu_p^{it}\right)^2 + \left(\Delta\mu_p^{it}\right)^2\right).$$

The following results holds for any (real-valued) positive martingale μ^t (see Neveu (1975), p.186):

$$\Pr\left(\sum_{t \geq 1}\left(\mu^t - \mu^{t-1}\right)^2 > a^2\right) \leq 3a^{-1}\mu^0.$$

This inequality applies to μ_p^{it} and μ_q^{it} which are martingales conditional on agent i's beliefs. Therefore

$$\Pr\left(\sum_{t \geq 1}\left(\Delta\mu_p^{it}\right)^2 > a^2 \middle| \mu^{i0}\right) \leq 3a^{-1}\mu_p^{i0}$$

But $\Pr\left(\sum_{t \geq 1}\left(\Delta\mu_p^{it}\right)^2 > a^2 \middle| \mu^{i0}\right) \geq \mu_s^{i0}\Pr\left(\sum_{t \geq 1}\left(\Delta\mu_p^{it}\right)^2 > a^2 \middle| s\right)$ so that

$$\Pr\left(\sum_{t \geq 1}\left(\Delta\mu_p^{it}\right)^2 > a^2 \middle| s\right) \leq 3a^{-1}\frac{\mu_p^{i0}}{\mu_s^{i0}} \quad \text{and} \quad \Pr\left(\sum_{t \geq 1}\left(\Delta\mu_q^{it}\right)^2 > a^2 \middle| s\right) \leq 3a^{-1}\frac{\mu_p^{i0}}{\mu_s^{i0}}$$

implying $\Pr\left(\sum_{t\geq 1}\left(\Delta v^{it}\right)^2 > 4a^2 \bigg| s\right) < \dfrac{3a^{-1}\left(\mu_p^{i0}+\mu_q^{i0}\right)}{\mu_s^{i0}}.$

This provides the result: if $\eta = 4a^2$ then defining the set $V_1(\eta,\varepsilon)$ by

$$\begin{cases} \mu_p^{i0} + \mu_q^{i0} < \dfrac{\mu_s^{i*}\varepsilon\eta^{1/2}}{12} \\[2mm] \mu_s^{i0} - \mu_s^{i*} < \dfrac{1}{2}\mu_s^{i*} \end{cases}$$

implies that $\Pr\left(\sum_{t\geq 1}\left(\Delta v^{it}\right)^2 > \eta \bigg| s\right) < \varepsilon$ if (μ^{10},μ^{20}) belongs to $V_1(\eta,\varepsilon)$.

Result 4: *Let us fix some $\eta,\varepsilon > 0$. There exists a neighborhood $V_2(\eta,\varepsilon)$ of $(\mu*^1,\mu*^2)$ such that if (μ^{10},μ^{20}) belongs to $V_2(\eta,\varepsilon)$ then for any of the two states of the world z not belonging to L (z=p or z=q),* $\Pr\left(\underset{t>0}{Max}\left(\mu_z^{it}\right) > \eta \bigg| s\right) \leq \varepsilon$

Proof: the "maximal inequality" for positive martingales (see Neveu (1975), p. 23) implies that:

$$\Pr\left(\underset{t>0}{Max}\left(\mu_z^{it}\right) > \eta \bigg| \mu^{i0}\right) \leq \dfrac{\mu_z^{i0}}{\eta}$$

and the inequality $\Pr\left(\underset{t>0}{Max}\left(\mu_z^{it}\right) > \eta \bigg| \mu^{i0}\right) \geq \mu_s^{i0}\Pr\left(\underset{t>0}{Max}\left(\mu_z^{it}\right) > \eta \bigg| s\right)$

leads to $\Pr\left(\underset{t>0}{Max}\left(\mu_z^{it}\right) > \eta \bigg| s\right) \leq \dfrac{\mu_z^{i0}}{\mu_s^{io}\eta}$

Therefore defining the set $V_2(\eta,\varepsilon)$ by the inequalities

$$\begin{cases} \mu_p^{i0} + \mu_q^{i0} < \dfrac{\varepsilon\eta}{2}\mu_s^{i*} \\[2mm] \mu_s^{i0} - \mu_s^{i*} < \dfrac{1}{2}\mu_s^{i*} \end{cases}$$

implies that $\Pr\left(\underset{t>0}{Max}\left(\mu_z^{it}\right) > \eta \middle| s \right) \leq \varepsilon$ if (μ^{10}, μ^{20}) belongs to $V_2(\eta,\varepsilon)$.

Result 5: *Let us fix some neighborhood W of $(\mu *^1, \mu *^2)$ in L^2 (that is, a set of "one-dimensional" belief pairs) and some $\varepsilon > 0$. There exists a neighborhood $V(\varepsilon)$ of $(\mu *^1, \mu *^2)$ in Δ^2 such that if (μ^{10}, μ^{20}) belongs to $V(\varepsilon)$ then*

$$\Pr\left(\underset{t\infty}{Lim}\left(\mu^{1t}, \mu^{2t}\right) \in W\right) > 1 - 3\varepsilon$$

Proof:

Notice first that $\left|\Delta u^{it}\right| = \dfrac{1}{2}\left|\Delta\mu_s^{it} - \Delta\mu_w^{it}\right|$ if s and w agree on one dimension and $\left|\Delta u^{it}\right| = 2^{-\frac{1}{2}}\left|\Delta\mu_s^{it} - \Delta\mu_w^{it}\right|$ otherwise so that in both cases $\left|\Delta u^{it}\right| \leq \dfrac{1}{2}\left|\Delta\mu_s^{it} - \Delta\mu_w^{it}\right|$.

Therefore the set of belief pairs

$$V'(\eta) = \left\{\mu_1, \mu_2 \middle| \mu_p^i < \eta;\ \mu_q^i < \eta;\ \left|u^i - u(\mu^{i*})\right| < (2K+1)\eta \text{ for } i = 1,2\right\}$$

converges towards $\left\{\mu *^1, \mu *^2)\right\}$ as η converges to zero.

Let us choose η such that $V'(\eta) \subset V_0$, $\overline{V'(\eta)} \cap L^2 \subset W$, and $\overline{V'(\eta)} \subset \left\{\left(\mu^1, \mu^2\right)\middle| \mu_w^1 > 0 \text{ and } \mu_w^2 > 0\right\}$.

Consider a pair of beliefs (μ^{10}, μ^{20}) belonging to

$$V_1(\eta,\varepsilon) \cap V_2(\eta,\varepsilon) \cap V'(\eta) \cap \left\{\left(\mu_1, \mu_2\middle\| u^i - u(\mu^{i*})\right| < \eta\right)\right\}$$

$(\mu^{10}, \mu^{20}) \in V_2(\eta, \varepsilon)$ implies by Result 4 that for i=1,2

$$\Pr\left(\underset{t>0}{Max}\left(\mu_p^{it}\right) < \eta \text{ and } \underset{t>0}{Max}\left(\mu_q^{it}\right) < \eta \middle| s \right) > 1 - 2\varepsilon .$$

$(\mu^{10}, \mu^{20}) \in V_2(\eta, \varepsilon)$ implies by Result 3 that for i=1,2

$$\Pr\left(\sum_{t \geq 1}\left(\Delta v^{it}\right)^2 > \eta \right) < \varepsilon$$

Therefore the event E:

$$\underset{t>0}{Max}\left(\mu_p^{it}\right) < \eta \text{ and } \underset{t>0}{Max}\left(\mu_q^{it}\right) < \eta \text{ and } \sum_{t \geq 1}\left(\Delta v^{it}\right)^2 > \eta$$

occurs with a probability greater than $(1 - 3\varepsilon)$.

Let us assume that E occurs, and let us assume that for some period t, and i=1 or i=2, $\left| u^{it} - u(\mu^{i*}) \right| > (2K+1)\eta$. Consider then

$$t_0 = Min\left\{ t \middle| \left| u^{1t} - u(\mu^{1*}) \right| > (2K+1)\eta \text{ or } \left| u^{2t} - u(\mu^{2*}) \right| > (2K+1)\eta \right\}.$$

The fact that E occurs implies than for all $t < t_0$, $(\mu^{1t}, \mu^{2t}) \in V'(\eta) \subset V_0$.

Therefore Result 2, applied to all pairs of beliefs (μ^{1t}, μ^{2t}) for $t < t_0$, implies that for i=1 and i=2

$$\left| u^{it_0} - u(\mu^{i*}) \right| \leq \left| u^{i0} - u(\mu^{i*}) \right| + \sum_{0 \leq t < t_0} \left| u^{it+1} - u^{it} \right|$$

$$\leq K\left(\sum_{0 \leq t < t_0}\left(v^{1t+1} - v^{1t}\right)^2 + \left(v^{2t+1} - v^{2t}\right)^2 \right) + \left| u^{i0} - u(\mu^{i*}) \right| \leq 2K\eta + \eta = (2K+1)\eta$$

Therefore it is impossible that $\left| u^{it_0} - u(\mu^{i*}) \right| > (2K+1)\eta$ for i=1 or i=2.

This implies that if (μ^{10}, μ^{20}) belonging to

$$\left(\mu^{10}, \mu^{20} \right) \in V_1(\eta, \varepsilon) \cap V_2(\eta, \varepsilon) \cap V'(\eta) \cap \left\{ \left(\mu_1, \mu_2 \left| u^i - u(\mu^{i*}) \right| < \eta \right) \right\}$$

then $\Pr\left(\left(\mu^{1t}, \mu^{2t} \right) \in V'(\eta) \text{ for all t} \right) > 1 - 3\varepsilon$

But (μ^{1t}, μ^{2t}) has to converge to some $(\mu^{1\infty}, \mu^{2\infty})$, and by Proposition 2 $\mu^{1\infty}$ and $\mu^{2\infty}$ belong to the same edge of Δ. We just showed that $\Pr\left(\left(\mu^{1\infty}, \mu^{2\infty} \right) \in \overline{V'(\eta)} \right) > 1 - 3\varepsilon$. Therefore the assumption that $\overline{V'(\eta)} \subset \left\{ \left(\mu^1, \mu^2 \right) \middle| \mu_w^1 > 0 \text{ and } \mu_w^2 > 0 \right\}$ implies that if $\left(\mu^{1\infty}, \mu^{2\infty} \right) \in \overline{V'(\eta)}$ then $\mu_w^{i\infty} > 0$ (i=1,2) so that $\left(\mu^{1\infty}, \mu^{2\infty} \right) \in L^2$. Therefore with a probability greater than $1 - 3\varepsilon$,

$$\left(\mu^{1\infty}, \mu^{2\infty} \right) \in \left(\overline{V'(\eta)} \cap L^2 \right) \subset W. \text{ \textbf{Q.e.d.}}$$

Proof of Proposition 5:

We consider a pair of initial beliefs $\left(\mu^{10}, \mu^{20} \right)$ such that for all pairs of states of the world (p,q)

$$\left| \frac{\mu_p^{10} \mu_q^{20}}{\mu_p^{20} \mu_q^{10}} - 1 \right| < \varepsilon.$$

We want to show that, if ε is small enough, then with probability 1. $\underset{t\infty}{Lim}\left(\mu^{1t}, \mu^{2t} \right) = \left(\delta_s, \delta_s \right)$.

Let us assume that with a positive probability

$$\underset{t\infty}{Lim}\left(\mu^{1t},\mu^{2t}\right)=\left(\mu^{1\infty},\mu^{2\infty}\right)\neq\left(\delta_s,\delta_s\right).$$

By proposition 2 we know that $\mu^{1\infty}$ and $\mu^{2\infty}$ belong to the same edge of Δ. Let us assume for example they both assign weight only to the states of the world (0,0) and (1,0). Let us define for i=1,2, $p_i=\mu^{i\infty}_{10}$, and let us assume that $p_1<p_2$ (group 1 is "to the left" of group 2).

Step 1:

Assume the signal direction is horizontal ($\theta=0$) and the signal intensity is maximal $(a=A)$.

The signal structure implies that for i=1,2:

$$B(\mu^{i\infty},S_{A,0})(1,0)=\frac{p_i(1+2A)}{p_i(1+2A)+(1-p_i)(1-2A)}=\frac{Cp_i}{Cp_i+1-p_i}\quad\text{and}$$

$$B(\mu^{i\infty},S_{A,\pi})(1,0)=\frac{p_i(1-2A)}{p_i(1-2A)+(1-p_i)(1+2A)}=\frac{p_i}{p_i+C(1-p_i)}.$$

With $C=\frac{1+2A}{1-2A}>1$

Since $\left(\mu^{1\infty},\mu^{2\infty}\right)$ is a steady state, it must be the case that whenever the speaker is in group 1, he wants to report the "left-wing signal" $S_{A,\pi}$ even after observing the "right-wing signal" $S_{A,0}$, which is true if

$$\left|\frac{p_1C}{p_1C+1-p_1}-\frac{p_2}{p_2+C(1-p_2)}\right|<\left|\frac{p_1C}{p_1C+1-p_1}-\frac{p_2C}{p_2C+1-p_2}\right|.$$

or $\left|\frac{p_1(1-p_2)C^2-p_2(1-p_1)}{p_2+C(1-p_2)}\right|<\left|\frac{(p_1-p_2)C}{p_2C+1-p_2}\right|$

which implies $\left|(p_1-p_2)C\right|>\frac{p_2C+1-p_2}{p_2+C(1-p_2)}\left|p_1(1-p_2)C^2-p_2(1-p_1)\right|.$

$C > 1$ and $p_1 < p_2$ imply that

$$\left|(p_1 - p_2)C\right| > \frac{1}{C}\left|p_1(1 - p_2)C^2 - p_2(1 - p_1)\right|, \text{ or equivalently}$$

$$C^2\left|1 - \frac{p_2(1 - p_1)}{p_1(1 - p_2)}\right| > \left|C^2 - \frac{p_2(1 - p_1)}{p_1(1 - p_2)}\right|.$$

Step 2:

For every period t, Bayes' rule implies that

$$\frac{\mu_{10}^{1t}\mu_{00}^{10}}{\mu_{00}^{1t}\mu_{10}^{10}} = \frac{\mu_{10}^{2t}\mu_{00}^{20}}{\mu_{00}^{1t}\mu_{10}^{10}} = \frac{\Pr(\text{messages up to } t|(1,0))}{\Pr(\text{messages up to } t|(0,0))}$$

so that $\dfrac{\mu_{10}^{1t}\mu_{00}^{2t}}{\mu_{00}^{1t}\mu_{10}^{2t}} = \dfrac{\mu_{10}^{10}\mu_{00}^{20}}{\mu_{00}^{10}\mu_{10}^{20}}$.

Taking the limit of this identity as t tends to infinity yields

$$\frac{p_2(1 - p_1)}{p_1(1 - p_2)} = \frac{\mu_{10}^{20}\mu_{00}^{10}}{\mu_{00}^{20}\mu_{10}^{10}}.$$

Step 3:

Steps 2 and 3 imply that if for all pairs (p,q) of states of the world,

$$\left|\frac{\mu_p^{10}\mu_q^{20}}{\mu_p^{20}\mu_q^{10}} - 1\right| < \varepsilon$$

then $C^2\left|1 - \dfrac{p_2(1 - p_1)}{p_1(1 - p_2)}\right| > \left|C^2 - \dfrac{p_2(1 - p_1)}{p_1(1 - p_2)}\right|$ and $\left|1 - \dfrac{p_2(1 - p_1)}{p_1(1 - p_2)}\right| < \varepsilon$

so that $\varepsilon C^2 > C^2\left|1 - \dfrac{p_2(1 - p_1)}{p_1(1 - p_2)}\right| > \left|C^2 - \dfrac{p_2(1 - p_1)}{p_1(1 - p_2)}\right| > C^2 - 1 - \varepsilon$

which cannot be true if ε is close enough to zero, given that $C > 1$.

This proves that if ε is close enough to zero, then $\left(\mu^{1\infty}, \mu^{2\infty} \right)$ cannot be a steady-state characterized by a one-dimensional conflict, so that (by Proposition 2), $\left(\mu^{1\infty}, \mu^{2\infty} \right) = \left(\delta_S, \delta_S \right)$. **Q.e.d.**

References

Aghion, P., P. Bolton, C. Harris, and P. Jullien, 1991, "Optimal learning by experimentation", *Review of Economic Studies*, vol. 58, 621-654.

Arrow, K., 1963, *Social choice and individual values*, New York : Wiley.

Banerjee, A. and R. Somanathan, 1999, "A simple model of voice", mimeo.

Crawford, V. and J. Sobel, 1982, "Strategic Information Transmission", *Econometrica*, vol. 50, No 6, 1431-1451.

Hart, S., 1985, "Nonzero-Sum Two-Person Repeated Games with Incomplete Information", *Mathematics of Operations Research*, vol.10, No 1, 117-153.

Neveu, J., 1975, *Discrete-Parameter Martingales*, Oxford: North-Holland..

Piketty, T., 1995, "Social Mobility and Redistributive Politics", *Quarterly journal of Economics*, vol. 110, No. 3, 551-584.

Poole, K.T. and H. Rosenthal, 1991, "Patterns of congressional voting", *American Journal of Political Science*, vol. 35, No 1, 228-278.

Snyder, J., 1996, "Constituency preferences: California ballot propositions, 1974-1990", *Legislative Studies Quarterly*, vol. 21, No. 4, 463-487.

Date Due

Lib-26-67